



Record No.	Region	City	Street Name	House No.	Description
1	上海	浦东	晨晖路	千	Native characters
2	Shanghai	Pudong	Chenhui Rd	1001	Romanized using both translation and transliteration
3	Shanghai	Purdong	晨晖路	千	Mixed scripts

Figure 11: Converting Non-Latin Characters into Latin Alphabets

The difference between translation and transliteration lies in preserving meaning (translation) versus preserving pronunciation across different languages (transliteration). Transliteration is a form of translation and is the practice of converting a text from one script into another. More precisely, transliteration is the representation of the letters of one alphabet with the letters (or combination of letters) of another alphabet on a consistent one for one basis.

In DS 4.2, the Global Address Cleanse (GAC) is capable of automatically detecting the input language and enabling the Romanization process (using both translation and transliteration techniques) to correct foreign addresses in both native and Latin scripts.

Using the All World address directly in conjunction with Global Address Cleanse you can cleanse and validate a foreign address down to last-line components in Latin script only. Using country-specific address directories for China and Russia in conjunction with Global Address Cleanse, you can cleanse and validate a foreign address (Chinese and Russian) down to address-line components in both Latin and Native (Chinese Kanji and Russian Cyrillic) characters.

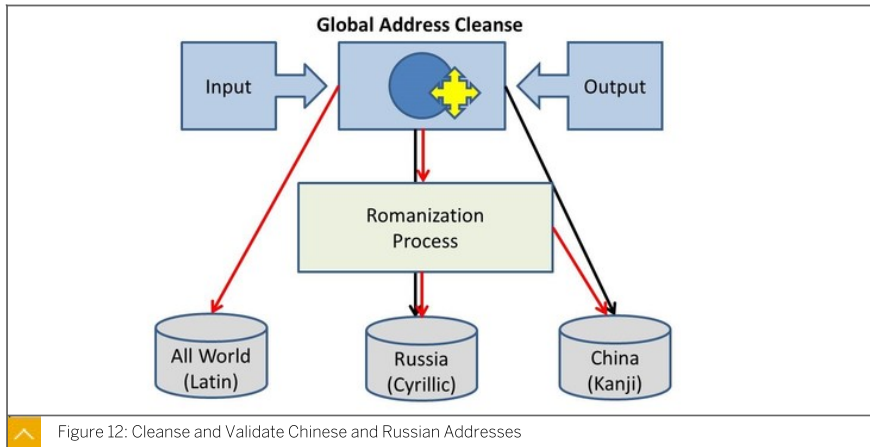


Figure 12: Cleanse and Validate Chinese and Russian Addresses

Often, a Western analysis needs to convert address data to English (Romanized) characters for duplication checks. In the figure Cleanse and Validate Chinese and Russian Addresses, the three records are identical, but written in different scripts. Without the knowledge of the local language, the only way to identify duplicate records in various scripts is to convert the address data into a single script.

In addition, Global Address Cleanse can also convert an address from native scripts to English (Latin) alphabets. In the Global Address Cleanse transform, there is an engine option within the Global Address Engine called "Output Address Script." If you select *Preserve*, GAC will